Control 1 MA-33A-1 2006-2

Profesor Dr. Gonzalo Hernandez - Auxiliar Gonzalo Rios

UChile - Departamento de Ingeniería Matemática

1) Representación Numerica

   (a) Considere una representación floating point tipo inicial, donde el número real más cercano a cero es 2^{-23}. Se sabe que la cantidad de bits de la mantisa es el doble que el número de bits usados en el exponente.

      i) Explique cuantos bits son necesarios para implementar esta codificación y como se distribuyen.

      ii) Sea $x = a_1 a_2 ... a_n$ la representación floating point de un número real, donde n es el números de bits utilizados por la codificación y:

$$a_1 = 1, \quad a_k = (a_{k-1} + k) \bmod 2, \; k = 2...n$$

      Determine el real representado por x.

      iii) Calcule el real máximo representable en esta codificación.

   (b) Considere la codificación floating point actual de 32 bits de la forma:

$$fl(x) = (-1)^s 2^z (1 + m)$$

   donde s el bit de signo, $z \in [-126, 127]$ el exponente y m la mantisa de 23 bits.

   Suponga que se redondea un número $x > 0$ para llegar a su codificación floating point.

      i) Demuestre que:

$$-2^{z-n} \leq x - fl(x) \leq 2^{z-n}$$

      donde n es la cantidad de bits de la mantisa.

      ii) Demuestre que $x \geq 2^z$ y que por lo tanto:

$$\frac{|x - fl(x)|}{x} \leq 2^{-n}$$

2) Aproximación y Errores

   (a) Determine la propagación de errores de las siguientes operaciones matemáticas:

i) $\phi(x, y) = 1 + \sin(x^2 + y^2), \quad x, y \in [0, 1]$

iii) $\phi(x, y, z) = \sqrt{xyz}$

Cuáles de ellas son estables/inestables ? Fundamente su respuesta.

(b) Considere las siguientes aproximaciones numéricas de la derivada de una función a variable real $f(x)$ en un punto x_0: (para $h \approx 0$)

$$\frac{df(x_0)}{dx} \approx \frac{f(x_0 + h) - f(x_0)}{h} \qquad \frac{df(x_0)}{dx} \approx \frac{f(x_0 + h) - f(x_0 - h)}{2h}$$

i) Haga un análisis de propagación de error de la primera fórmula con $\varepsilon_h = 0$.
Obs: Si aparece $\frac{f(x+h)-f(x)}{h}$, aproxímelo por $f'(x)$.

ii) Para una aritmética finita de 5 cifras significativas con redondeo, aproxime la derivada de $f(x) = x \ln x$ en el punto $x_0 = 10.125$, usando la segunda fórmula con $h = 0.001$. Compare este valor aproximado con el valor exacto de la derivada. Calcule el error absoluto y relativo.

3) Dada una matriz tridiagonal, su factorización $A = LU$ puede ser llevada a la forma:

$$\begin{bmatrix} a_{11} & a_{12} & 0 & \cdots & 0 \\ a_{21} & a_{22} & a_{23} & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn-1} & a_{nn} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots \\ 0 & 0 & l_{nn-1} & l_{nn} \end{bmatrix} \begin{bmatrix} 1 & u_{12} & 0 & \cdots & 0 \\ 0 & 1 & u_{23} & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & u_{n-1n} \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

mediante el Método de Crout:

Paso 1:	Paso 2: Para $i = 2, ..., n-1$	Paso 3:
$l_{11} = a_{11}$ $u_{12} = \dfrac{a_{12}}{l_{11}}$	$l_{i(i-1)} = a_{i(i-1)}$ $l_{ii} = a_{ii} - l_{i(i-1)}u_{(i-1)i}$ $u_{i(i+1)} = \dfrac{a_{i(i+1)}}{l_{ii}}$	$l_{n(n-1)} = a_{n(n-1)}$ $l_{nn} = a_{nn} - l_{n(n-1)}u_{(n-1)n}$

(a) Calcule la factorización de Crout de la siguiente matriz:

$$A = \begin{bmatrix} 5 & -2 & 0 & 0 \\ -2 & 5 & -2 & 0 \\ 0 & -2 & 5 & -2 \\ 0 & 0 & -2 & 5 \end{bmatrix}$$

(b) Calcule la cantidad de *ops* que este método realiza aplicado a una matriz cuadrada tridiagonal de tamaño n.